DOCUMENT RESUME

ED 172 201 CS 204 846

AUTHOR	Friedman, Myron; Fowler, Elaine
TITLE	Assessing Elementary Students' Writing Skills. Publication No. 78.74.
INSTITUTION	Austin Independent School District, Tex. Office of Research and Evaluation.
PUB DATE	Apr 79
NOTE	23p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 8-12, 1979)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Composition (Literary); Elementary Education; Elementary School Students; *Evaluation Methods; Measurement Instruments; *Program Evaluation; Student Development; *Student Evaluation; *Writing Skills

ABSTRACT

An instrument was developed for use in the evaluation of a pilot program to improve the writing skills of elementary school students in the Austin (Texas) Independent School District. Called the "Assessment of Writing Skills" (AWS), the instrument assesses writing maturity, productivity, and writing mechanics by collecting a holistic evaluation and such frequency counts as total words, percentage of words spelled correctly, total sentences, sentence length, total correct punctuations, percentage of correct punctuations, total correct capitalizations, paragraph usage, number of modifying words, and sentence usage. The AWS was used by three raters to evaluate 268 second through fifth grade students' descriptive writing samples during the fall (pretest) and spring (posttest) of the same academic year. A measure of interscorer/intrascorer reliability ensured that the essays were rated consistently. The data reveal that fourth graders had lower pretest scores than third graders on eight of 13 variables, a fact that might be explained by the corresponding shift from printing to cursive writing during that period. The results also suggest that the AWS meets criteria for the evaluation of instructional programs; the holistic scores of AWS can assess overall growth over time, while frequency counts can provide feedback regarding specific skill development and needs for instruction. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED172201

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ASSESSING ELEMENTARY STUDENTS' WRITING SKILLS

A paper presented at the annual meeting of the
American Educational Research Association

April, 1979

Authors: Myron Friedman, Austin I.S.D.
Elaine Fowler, University of Texas at Austin

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Myron Friedman
Elaine Fowler
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Office of Research and Evaluation
Austin Independent School District

Austin, Texas

Pub. No.: 78.74

CS 204 846

ASSESSING ELEMENTARY STUDENTS'WRITING SKILLS

The report of the National Assessment of Educational Progress (Writing Mechanics, 1975) emphasized that there is work to be done in improving children's writing at all levels. Specifically, the report indicated that nine-year olds had not mastered the basic skills of writing, that only the best 13-year olds had, and that by age 17, only 50% of the teenagers could put together simple sentences, use commas, and express simple ideas (Wolfe, 1979).

The teaching of writing has always been part of formal schooling, yet many educators realize that they themselves know little about the act of writing and even less about ways of evaluating children's writing and helping them to improve their writing skills. For the most part, schools only provide opportunities for children to write, rather than opportunities to learn how to write better.

Writing has received a considerable amount of attention in the last few years, however, in both popular and academic journals. At the same time, writing process research has brought new information about what is involved in writing and how it ought to be taught. However, much remains to be determined about the evaluation and measurement of children's writing.

ALTERNATIVES FOR ASSESSING WRITING

Considerable attention over measuring improvement in student writing has focused on the question of what type of measure to use. Two different types of measurement are the objective test and the essay. The typical objective test can be standardized and made reliable, yet it suffers from low face validity because it only indirectly measures writing ability. The essay, which measures longer units of discourse and is more content-

oriented, has higher face validity since it is an actual sample of student writing, but it entails problems of writer fluctuation and reader reliability. In an attempt to resolve the controversy, recent research in the improvement of student writing has recommended the use of essays utilizing rating scales in conjunction with objective testing of some type.

Essays can be assessed in a manner which provides the high face validity required by educators as well as the specificity and high reliability required by researchers. This can be accomplished by means of a rating strategy which includes Frequency Counts as well as a holistic assessment of the essays.

FREQUENCY COUNTS

Frequency counts have become an important type of objective measurement. The importance of the frequency count (in contrast to rating procedures for an essay) lies in its potential for describing a composition in fairly objective terms which can have meaning to both researchers and teachers. Frequency counts provide knowledge about particular features of a language which are important to many of us. We often teach in terms of the specific and certainly our ability to talk about writing depends on the existence of reliably defined features and categories.

HOLISTIC RATING

Holistic evaluation is based on the idea that a valid test of discourse depends upon the examination of a sample of discourse as a whole, not merely the collection of its parts (Lloyd-Jones, 1977). Although holistic evaluation does not result in the sort of detailed descriptions needed for diagnosing students' writing problems, it does get us closer to what is essential in a piece of writing--communication as a whole message. We are defining

holistic evaluation here as Cooper does (1977) to mean any procedure which stops short of enumerating linguistic, rhetorical or informational features of a piece of writing. Some holistic procedures may specify a number of particular features and even require that each feature be scored separately, but the reader is never required to stop and count or tally incidents of the features. The reader uses a list of features only as a general guide in reaching a holistic judgment (Cooper, 1977).

An important principle of holistic evaluation is the understanding that the score assigned to a given essay is meaningful only in relation to the other essays comprising the group in which it was rated and not in relation to external standards.

The inter-rater reliability correlations (measure of the extent to which raters agree with one another on the rating assigned to a given essay) reach as high as .70 to .80. The reliability becomes higher as raters are given special training prior to work.

THE STUDY

Considering the relative merits of frequency counts and holistic ratings, an instrument was developed for use in the evaluation of an ESAA Pilot Program designed to improve the writing skills of low SES elementary-grade students (Friedman, 1978). The writing instruction program was implemented in grades K-5 in six elementary schools in the Austin Independent School District. The program was designed to improve students' writing skills in the areas of writing mechanics and grammar as well as increase the amount which students wrote. The writing assessment instrument therefore was designed to assess writing maturity, productivity, and writing mechanics. The instrument will be referred to as the Assessment of Writing Skills (AWS).

Thirteen different aspects of writing were identified for inclusion

in the AWS. These included the total number of words, percent of words spelled correctly, total number of sentences, sentence length, total number of correct punctuations, percent correct punctuations, total number correct capitalizations, paragraph usage, number of modifying words, and sentence usage. These aspects of writing were to be measured by frequency counts. In addition, a holistic rating of overall quality was included.

Procedure: The instrument was used by three raters to assess the writing skills of 268 students in grades two through five. The assessment was based upon samples of descriptive writing produced by the students in the fall and spring of the same academic year. At both time periods, the students wrote in response to an identical stimulus. Attachment 1 is a description of the procedures used in generating the writing samples.

Rating of Papers: The second author developed the AWS and provided three raters with a half-day training session in the procedures for rating writing samples. Attachment 2 includes the form developed (the AWS) and directions for its use.

During the training session the raters each scored the same essays, discussed the scores they marked, and deliberated over any discrepancies about their ratings. After obtaining a high level of interscorer reliability, each rater proceeded to independently rate 16 folders of student essays. The essays rated by each rater were selected according to the following criteria:

(1) Each student's pre- and posttest essay would be rated during the same rating session by the same rater.

(2) During each rating session, each rater would rate a set of essays which was balanced to include pre- and posttest essays.

Over the three-day period, each rater assessed 90 essays.

Reliability of Ratings: In order to obtain measures of interscorer reliability (extent to which all three scorers rated essays the same) and intrascorer reliability (extent to which the same scorer rated essays the same), a number of essays from each treatment group and grade were Xeroxed and distributed in such a manner that each rater scored some essays twice and some essays were scored by all three raters.

Scores on these duplicate essays were used to compute the measures of interscorer and intrascorer reliability. The reliability coefficients were computed by means of the INTRAR program of the VSTAT library. The interrater reliability on the 11 items of the Writing Assessment Instrument are listed on Table 1. For items one through seven and item nine, the reliability of the raters exceeds .9 (1.00 indicates perfect reliability). The item having the lowest interrater reliability was item 11 at .78. These high reliability ratings are a good indication that each of the three raters scored the essays in a similar manner. An essay would have received a similar score regardless of who the rater was.

TABLE 1

INTERRATER AND INTRARATER RELIABILITIES
ON EACH OF 11 ITEMS ON THE
ASSESSMENT OF WRITING SKILLS

ITEM	INTERRATER RELIABILITY	INTRARATER RELIABILITIES		
		RATER 1	RATER 2	RATER 3
TOTAL WORDS	.9985	.9979	.9998	1.0000
WORDS SPELLED WRONG	.9386	.9255	.9632	.9672
TOTAL SENTENCES	.9486	.8727	1.0000	.9900
TOTAL CORRECT PUNCTUATION	.9076	.8054	.9970	.9934
NUMBER PUNCTUATION NEEDED	.9600	.7213	.9949	.9852
TOTAL CORRECT CAPITALS	.9714	.9938	.9949	.9208
NUMBER CAPITALS NEEDED	.9549	.9065	.9540	.8442
PARAGRAPH USAGE	.8728	.9114	.6203	1.0000
TOTAL MODIFYING WORDS	.9877	.9911	.9793	.9979
SENTENCE USAGE	.8580	.8814	.5000	.8792
OVERALL QUALITY	.7837	1.0000	.5745	.6842

The intrarater reliabilities for each rater are also listed on Table 1. In reviewing each rater's reliability on the 11-scale items, it is apparent that they were all highly reliable on most of the items (reliability was close to perfect). Each rater did have a slightly lower reliability on several items. Rater 1 had the lowest reliability on items 10 and 11, .50 and .57 respectively. Rater 2's lowest reliability was on item five (.72). Rater 3 had her lowest reliability on item 11 (.68). Overall, however, intrarater reliabilities were very high, ensuring that the essays were all rated in a consistent manner.

Analyses: In analyzing the AWS, 13 variables were computed from the original 11 rating items. This list of variables is found on Attachment 3. In order to examine the areas of writing measured by the AWS and to determine if it measures the same thing pre and post, a series of factor analyses were done. First a noniterative, principal component analysis with no rotation was done to determine the number of factors. Then principal component factor analyses with orthogonal and oblique rotations were done for the pre- and posttests. The factor analyses were done using the SPSS Factor Analysis program on the University of Texas at Austin's CDC 6000 computer.

The orthogonal rotation was a Verimax rotation with Kaiser normalization. The oblique rotations were direct oblimin rotations with the delta value set at 0.

Results:

Number of Factors: An inspection of the eigenvalues resulting from the principal components analyses showed that the AWS measured four factors at pretest and three factors at posttest.

Naming the Factors: Because the AWS appeared to measure the same factors at each grade level, all grades were combined for a principal axes solution with orthogonal and oblique rotations. The orthogonal rotation produced the most meaningful results. The loading order for the variables at pretest and posttest are found on Tables 2 and 3. Although the factor loadings were not identical for pretest and posttest, they were very similar. At pretest the factors identified were 1) Overall Quality--Productivity, 2) Overall Quality--Maturity, 3) Mechanics--Punctuation, and 4) Mechanics--Spelling and Capitalization. At posttest the factors identified were 1) Overall Quality--Productivity, 2) Overall Quality--Maturity, and 3) Mechanics--Punctuation and Capitalization. Overall Quality--Productivity includes the Total Number of Sentences and Total Number of Words written. The major aspect of this factor appears to be the fluency of writing. Overall Quality--Maturity includes Sentence Length, Total Modifying Words, Sentence Usage, and Paragraph Usage. The major aspect of this factor appears to be the syntactic complexity of the AWS rather than the quantity.

TABLE 2

LOADING ORDER OF THIRTEEN PRETEST WRITING ASSESSMENT VARIABLES ON ORTHOGONAL FACTORS

VARIABLES	ORTHOGONALLY ROTATED FACTORS			
	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
Total Correct Capitals	(.90)	.05	.15	.26
Total Number Sentences	(.86)	.10	.10	.15
Total Words	(.73)	.60	.11	.04
Sentence Length	-.04	(.65)	.14	.09
Total Modifying Words	.59	(.63)	.07	.27
Sentence Usage	.35	(.60)	.13	.15
Overall Quality	.50	(.37)	.20	.28
Paragraph Usage	.06	(.39)	-.02	-.02
Total Correct Punctuation	.32	.21	(.78)	-.03
Percent Correct Punct.	-.01	.02	(.72)	.37
Percent Spelled Correct	.19	.30	.09	(.41)
Percent Correct Capitals	.22	-.13	.19	(.52)
Percent Modifying Words	.09	.46	.03	(.58)

TABLE 3

LOADING ORDER OF THIRTEEN POSTTEST WRITING ASSESSMENT VARIABLES ON ORTHOGONAL FACTORS

VARIABLES	ORTHOGONALLY ROTATED FACTORS		
	FACTOR 1	FACTOR 2	FACTOR 3
Total Number Sentences	(.88)	-.06	.20
Total Modifying Words	(.83)	.42	.04
Total Correct Capitals	(.83)	.03	.37
Total Words	(.78)	.42	.12
Overall Quality	(.62)	.52	.15
Percent Modifying Words	(.37)	.21	.02
Sentence Length	-.02	(.63)	.14
Paragraph Usage	.21	(.57)	-.02
Sentence Usage	.43	(.54)	-.00
Percent Spelled Correct	.29	(.33)	.27
Percent Correct Punctuation	-.02	.05	(.85)
Total Correct Punctuation	.16	.21	(.60)
Percent Correct Capitals	.14	-.03	(.46)

This analysis suggests that the holistic rating (Overall Quality) actually includes two separate aspects of writing: Productivity and Maturity. This finding is supported by an examination of the relationship between the holistic rating of Overall Quality and each of the other 12 variables. These correlations are found on Table 4. There is a moderately strong to weak relationship between Overall Quality and each of the other items. This finding suggests that the Overall Quality rating may be a powerful indicator of the relative quality of the writing sample; however, it does not include a full consideration of the more specific aspects of writing, nor does it adequately replace them.

Additional Findings: In using the AWS to assess the effectiveness of the ESAA writing program, the following observations were made:

TABLE 4

CORRELATION OF PRETEST-POSTTEST HOLISTIC RATING (OVERALL QUALITY) WITH FREQUENCY COUNTS ON WRITING ASSESSMENT FOR SECOND, THIRD, FOURTH, AND FIFTH GRADERS

VARIABLE	GRADE 2 (N=88)	GRADE 3 (N=54)	GRADE 4 (N=58)	GRADE 5 (N=68)
Total Words				
Pretest	.48	.30	.57	.77
Posttest	.50	.52	.51	.63
Percent Spelled Correct				
Pretest	.29	.40	.51	.28
Posttest	.41	.35	.33	.32
Total Number of Sentences				
Pretest	.65	.38	.51	.49
Posttest	.40	.51	.41	.55
Sentence Length				
Pretest	.28	.24	.24	.18
Posttest	.36	.49	.12	.05
Total Correct Punctuation				
Pretest	.62	.31	.57	.19
Posttest	.36	.17	.16	.22
Percent Correct Punctuation				
Pretest	.38	.18	.46	.11
Posttest	.31	.07	.23	.02
Total Correct Capitals				
Pretest	.44	.33	.49	.53
Posttest	.49	.44	.44	.59
Percent Correct Capitals				
Pretest	.18	.08	.48	.08
Posttest	.20	.33	.29	.01
Paragraph Usage				
Pretest	.26	.19	.11	.41
Posttest	.17	.33	.48	.33
Total Modifying Words				
Pretest	.65	.35	.54	.77
Posttest	.53	.50	.61	.69
Percent Modifying Words				
Pretest	.43	.27	.39	.37
Posttest	.46	.40	.53	.30
Sentence Usage	.51	.26	.56	.65
Pretest	.52	.57	.55	.47
Posttest				

- Significant pretest to posttest gains were found for second graders on 11 of 13 variables; for third graders on only one variable; for fourth graders on six of 13 variables; and for fifth graders on only two variables.
- Fourth graders had lower pretest scores than did third graders on eight of the 13 variables. These findings are illustrated on Table 5, which lists the pre- and posttest means and standard deviations for each grade on the 13 variables of the AWS.

Inspection of program implementation data and student achievement data did not provide an explanation for these findings. However, a possible explanation for these findings may be the switch from printing to cursive writing. Third graders in the study were printing at pretest and were able to demonstrate skills developed during the second grade. At posttest; however, they were writing in cursive script. Attempts at mastering this new writing skill may have hampered their ability to achieve significant gains in the skills being measured by the AWS. Fourth graders at pretest may have still been at a disadvantage in attempts to display their skills, cursive writing was still a new experience for them. By posttest; however, they may have adequately overcome this handicap.

IMPLICATIONS FOR PROGRAM EVALUATION

Hammil and Larsen (1978) have recommended that the purpose for assessing writing be made explicit before selecting an assessment strategy. They have identified four possible purposes: 1) To identify students needing remediation, 2) to determine students' strengths and weaknesses, 3) to document progress, and 4) for use in research and evaluation. Depending upon the purpose, different assessment strategies will be appropriate. This summary

TABLE 5

PRE- AND POSTTEST MEANS AND STANDARD DEVIATIONS
FOR EACH GRADE ON THE 13 VARIABLES OF THE AWS

VARIABLE	GRADE							
	SECOND		THIRD		FOURTH		FIFTH	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
Total Words								
Pretest	24.92	16.54	35.35	20.87	34.43	18.15	61.46	33.75
Posttest	**35.32	20.44	41.06	19.47	**45.93	18.47	68.13	37.29
Percent Spelled Correct								
Pretest	.69	.24	.83	.14	.84	.16	.86	.14
Posttest	**.77	.19	.85	.16	.87	.09	.89	.09
Total Number Sentences								
Pretest	4.07	3.17	5.46	3.50	4.93	3.20	7.01	4.36
Posttest	**5.13	3.38	5.31	2.94	5.47	3.03	7.12	4.03
Sentence Length								
Pretest	4.96	3.55	6.76	3.32	7.26	5.55	9.80	6.41
Posttest	**6.44	3.92	7.19	4.18	8.40	4.88	11.19	8.96
Total Correct Punc.								
Pretest	2.05	2.61	4.19	3.83	3.78	4.42	5.79	7.00
Posttest	**3.41	3.32	5.30	5.11	*5.60	7.42	6.19	7.38
Percent Correct Punc.								
Pretest	.42	.52	.59	.46	.53	.43	.51	.33
Posttest	.50	.41	.61	.35	.58	.34	.51	.39
Total Correct Capitals								
Pretest	4.10	3.08	6.26	.55	5.12	4.37	8.71	5.33
Posttest	**5.50	3.55	5.83	.39	6.29	3.34	8.16	5.15
Percent Correct Capitals								
Pretest	.70	.38	.88	.17	.76	.37	.84	.23
Posttest	*.85	.44	.81	.25	.81	.24	.80	.43
Paragraph Usage								
Pretest	.17	.38	.26	.56	.40	1.09	.51	.53
Posttest	.23	.42	.35	.52	.50	.57	*.79	.56
Total Modifying Words								
Pretest	6.18	6.34	9.65	7.36	9.43	6.67	22.18	14.72
Posttest	**11.02	8.90	*12.48	8.82	**15.07	8.96	25.10	15.33
Percent Modifying Words								
Pretest	.22	.16	.26	.11	.26	.12	.34	.11
Posttest	**.27	.15	.27	.13	*.31	.12	.39	.34
Sentence Usage								
Pretest	.94	.59	1.31	.64	1.17	.63	1.66	.80
Posttest	**1.22	.73	1.33	.75	*1.47	.82	*1.94	.73
Overall Quality								
Pretest	1.57	.52	1.91	.40	1.95	.48	2.46	.76
Posttest	**1.81	.58	1.96	.51	*2.16	.62	2.60	.67

** $p<.01$
* $p<.05$

will address the use of assessment for evaluation purposes.

For use in evaluating an instructional program, an assessment instrument should be designed to provide answers to the following questions:

1. Was there any growth in writing skills as a result of the program?
2. What aspects of writing were most/least effected by the program?

In order to answer these questions, the instrument should have the following qualities:

1. It should have high face validity; it should appear to be measuring those aspects of writing which the instructional program have impact upon as well as reflect general consensus among teachers as to what good writing entails.
2. It should be readily useable, requiring only a minimal amount of training.
3. It should provide relatively high measures of inter- and intrarater reliabilities.
4. It should provide a holistic rating of the students' writing as well as an indication of the students' performance in specific areas such as writing mechanics and grammar.

The results of the present study suggest that an instrument such as the AWS, which combines frequency counts with a holistic score does meet the above criteria. The holistic score can be used to assess the overall growth over time, while frequency counts provide feedback regarding specific skill development and needs for instruction.

I. Ask students to clear their desks. Prepare students for experience by means of an introductory activity which will sensitize them to their five senses.

A. Show the class a picture for a few seconds, then remove it from sight. Ask the children to describe the things which they saw in the picture. In your discussion emphasize that we all note different things when using our sense of sight.

B. Have the children sit absolutely quiet for 15 seconds. Ask them to describe the things they heard while everyone was quiet.

C. Have the students describe what things they can feel while sitting still in their seats.

D. Ask the students what things they might smell in the cafeteria at lunch time. Ask them to describe those smells

E. Ask the students to explain how different things taste to them. (eg. lemon, orange, chocolate, onions, hamburgers)

II. Tell the students that we will be going on a walk through different parts of the school. During the walk they must be absolutely quiet, anyone who cannot will not be able to participate. As we go on our walk, pay attention to what you see, hear, smell, touch. Try to describe these different things to yourself while we are on the walk. We will make several stops along the way in order to describe the things which we experience. We won't have time for everyone to speak; please don't talk unless you are called upon. When we come back to the classroom everyone will have a chance to write about the things you experienced. When we come back to the classroom, quickly go back to your seat and remain quiet for further instructions.

III. Go on the walk.

A. Go through the halls into the library, office, and cafeteria.

B. Stop for a moment after you leave each of the above locations. Ask the students to quickly describe some of the things they saw, heard, etc. If the students do not respond, then quickly summarize for them the things which your experienced.

C. Reenter the classroom.

IV. Handout paper and pencils. Ask students to write their first name (plus last initial).

A. Ask the students to sit quietly for a short while (30 seconds) and think about the walk and all of the things which they saw, heard, etc.

B. Ask the students to write about their experience as if explaining it to a friend. Write explaining/describing what you saw so that your friend will have a written picture of your experience. Tell the students that they will have up to 15 minutes for the writing.

ATTACHMENT 1: PROCEDURE FOR OBTAINING WRITING SAMPLE

Rater's Name: _____

AUSTIN INDEPENDENT SCHOOL DISTRICT
Office of Research and Evaluation

Day: _____

Time: _____

ASSESSMENT OF WRITING SKILLS

Pupil No. _____

_____ 1. Total number of words written.

_____ 2. Total number of words spelled incorrectly.

_____ 3. Total number of sentences pupil wrote. (capitalized or punctuated correctly or not)

_____ 4. Number of places pupil punctuated correctly.

_____ 5. Total number of places where punctuation was needed.

_____ 6. Number of words child capitalized correctly.

_____ 7. Total number of words needing capital letters.

_____ 8. Paragraphs
0 if pupil listed words, phrases, or sentences.
1 if pupil used paragraph form.
2 if pupil had two paragraphs or more.

_____ 9. Total number of modifying words. (adjectives, adverbs, noun markers, prepositional phrases)

_____ 10. Types of sentences
3 if compound, complex and simple sentences are used.
2 if compound and simple or complex and simple, etc. are used.
1 if simple or only one type is used.
0 if no sentences, only words or phrases are used.

11. Overall Quality of Composition

Circle One.

A B C D F

ATTACHMENT 2: ASSESSMENT OF WRITING SKILLS AND DIRECTIONS FOR ITS USE (Page 1 of 5).

DIRECTIONS FOR USING ASSESSMENT OF WRITING SKILLS

Items 1-10 on the rating scale are counting exercises. The following directions should be used for each composition you rate. You may mark on the composition for marking and counting purposes.

1. Count the total number of words and record the number of words on the first line of the scale. Give credit if you can phonetically discern something resembling a word. If two words are put together as one; give credit for two words if that is how the words are spelled in standard English (eg., teacup = 1 word, tankyo = 2 words). If the pupil wrote one letter or two, and not the complete word, don't give credit.

2. Next circle the number of misspelled words; count them and enter the number on line two of the scale. Count contractions wrong if no apostrophe.

3. Put slash marks at the end of every sentence even if incorrectly punctuated. Use your judgment for distinguishing one sentence from another. Give credit for a sentence(s) that is joined by <u>and</u> (a run-on) if you think what follows <u>and</u> makes sense. If what follows the <u>and</u> does not continue the topic or thought of the sentence put a slash mark (/) to indicate that you think a new sentence should begin there. Then count the number of slash marks. Enter the number of slashes (sentences) in line 3 of the scale.

4. Count the number of punctuation marks (.,?!;:" ") the pupil used correctly and put on line 4. If he had punctuation marks that were used incorrectly, don't give credit for them.

5. Now go through the paper and count the punctuation marks that were needed in the composition (including those that the pupil already had correct) and enter on line 5. One point should be tallied for each part of a quotation mark ("<u> </u>" = 2).

6. Count the number of words the pupil capitalized correctly and put the tally on line 6. Underline those words needing capital letters for counting purposes.

7. Now go back through the composition and underline and count the number of words that needed capital letters (including those the pupil capitalized correctly) and record on line 7.

8. If the paper listed words, phrases, or sentences--each starting a new line--put a 0 in the blank. If the pupil used paragraph form (whether indented or not)put a 1 in the blank. If two or more paragraphs were used, put a 2 in the blank.

9. Put brackets [ ] around every noun marker, adjective, adverb or prepositional phrase, you see. If there is a noun marker inside the prepositional phrase--for example [in [the] wall]--, give them 2 points --one for the phrase and one for the marker. Count the number of brackets and enter this number on line 9.

10. Put ( ) parentheses around any dependent (subordinate) or independent (main) clauses you see. Then look over the composition to count the types of sentences you see. If the sentences are all simple score 1. If you see two types, simple and compound or simple and complex, put a 2 in the blank. If the child used all three types, put a 3 in the blank. Don't count the number of sentences just the types you see. Your number will never exceed 3 for item 10. If there were no complete sentences, put a 0 in the blank. Here are some words that usually initiate a dependent/subordinate clause:

when we went	whom
where she lived	unless
that he could win	because
who buy paintings	so that
which he fought	as long as

11. For item number 11 on the scale, read the following instructions.

For this rating, you are asked to use the following 5 point scale.

A	Good
B	Above Average
C	Average
D	Poor
E	Bad

For purposes of a rating session such as we are doing--in which amounts of improvement will be compared--not grading papers to return to students, it is very important that a spread of grades/ratings be recorded. If the grades/ratings are bunched at the C--level, then it will be impossible to detect differences in amount of improvement. Therefore, you are urged to use the whole range of grades A-F as much as possible. Try to be sensitive to even relatively minor differences in quality and do not "play it safe" by using the C category more than necessary. A special caution must be observed regarding

the impression created by handwriting or general appearance of the paper. Previous research has indicated that handwriting and appearance are factors in determining the rating a paper receives. Ideally all compositions should be typed in standard form with all spelling and other errors preserved. Since it was not possible to have these compositions typed, it will be up to you, the reader, to discount handwriting and appearance as much as possible.

Read through the criteria listed for an A,B,C,D,F and then reread the composition and rate it on item 11.

Criteria for Item 11.

Ranking of A

a. Wealth of ideas and details.
b. Colorful words.
c. Combination of clauses and sentence types.
d. Correct use of punctuation, capitalization, and spelling most of the time.
e. Cause and effect developed.
f. Many modifiers used including phrases.
g. Quite a few words written in composition.

Ranking of B

a. Ideas clear—many/some details.
b. Sporadic use of vivid words.
c. Use of longer sentences. Sentences start in varied patterns (not just subject-verb object).
d. Has modifiers used in several places.
e. Not as many phrases as single word modifiers.
f. Quite a few words in composition—various sentence types, at least 2.
g. Some inappropriate use of capitalization, punctuation, and spelling.

Ranking of C

a. Some descriptive words—varied adjectives, etc.
b. Some inappropriate use of capitalization and punctuation.
c. Some words spelled incorrectly.
d. Ideas clear but lack full development—few details.
e. Uses noun markers and some simple modifiers.
f. More than just a few words written.

Ranking of D

a. Some specific verbs, nouns, adjectives, adverbs.
b. Inadequate details--stilted, restrained ideas which prevent smooth relationships.
c. Use of short, simple sentences, subject verb--object order. Sentences start the same way--"I saw.....".
d. Uses some noun markers.
e. Not many words written.
f. Composition makes sense, but dull.

Ranking of F

a. Limited repertoire of words--very few.
b. No details, bareness of expression.
c. Use of short simple sentences that start the same way or incomplete sentences.
d. Little or no use of punctuation or capitalization.
e. Few or no noun markers, modifiers.
f. Composition doesn't make much sense to the reader.

Variable 1: TOTAL NUMBER OF WORDS

Item 1 from rating form. Blank if item 1 is blank.

Variable 2: PERCENT SPELLED CORRECTLY

(Item 1 - Item 2)/Item 1. Blank if item 1 is blank or zero, if item 2 is blank, or if item 2 is greater than item 1.

Variable 3: TOTAL NUMBER OF SENTENCES

Item 3 from rating form. Blank if item 3 is blank or item 1 is zero.

Variable 4: SENTENCE LENGTH

Item 1/Item 3. Blank if item 1 is blank or zero, or if item 3 is blank. Equal to item 1 of item 3 is 0.

Variable 5: TOTAL NUMBER CORRECT PUNCTUATIONS

Item 4 from rating form. Blank if item 4 is blank or item 1 is zero.

Variable 6: PERCENT CORRECT PUNCTUATIONS

Item 4/Item 5. Blank if item 4 or item 5 is blank, item 1 is zero, or item 4 is greater than item 5. Equal to 1.0 if item 4 and item 5 are both zero.

Variable 7: TOTAL NUMBER CORRECT CAPITALIZATIONS

Item 6 from rating form. Blank if item 6 is blank or item 1 is zero.

Variable 8: PERCENT CORRECT CAPITALIZATIONS

Item 6/Item 7. Blank if item 6 or item 7 is blank, item 1 is zero, or item 6 is greater than item 7. Equal to 1.0 if item 4 and item 5 are both zero.

Variable 9: PARAGRAPH USAGE.

Item 8 from rating form. Blank if tiem 8 is blank or item 1 is zero.

ATTACHMENT 3: COMPUTATION OF VARIABLES. DEFINITIONS AND SPECIFICATIONS USED TO CALCULATE VARIABLES ANALYZED. (PAGE 1 OF 2)

Variable 10: TOTAL NUMBER OF MODIFYING WORDS

Item 9 from rating form. Blank if item 9 is blank or item 1 is zero.

Variable 11: PERCENT MODIFYING WORDS

Item 9/Item 1. Blank if item 9 is blank or item 1 is blank or zero.

Variable 12: SENTENCE USAGE

Item 10 from rating form. Blank if item 10 is blank or item 1 is zero.

Variable 13: OVERALL QUALITY

Item 11 from rating form. Blank if item 11 is blank.

REFERENCES

Cooper, C. and Odel, L. Evaluating writing: Describing, measuring judging. Urbana: National Council of Teachers of English, 1977.

Friedman, M. H. and Matuszek, P. ESAA pilot 1977-78, final technical report. Austin, Texas: Austin Independent School District, Office of Research and Evaluation, June, 1978.

Hammill, D. D. and Larsen, S. C. The test of written language. Austin: Pro-Ed, 1978.

Lloye-Jones, R. Primary trait scoring. In Cooper, C. and Odell, L. Evaluating writing: Describing, measuring, judging. Urbana: National Council of Teachers of English, 1977.

Wolfe, D. T. Back to the children--the real basics. Language Arts, February, 1979, 146.